Title: Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Authors – Sasha R Howard,[1] Leonardo Guasti,[1] Ariel Poliandri,[1] Alessia David,[4] Claudia P Cabrera,[2,3] Michael R Barnes,[2,3] Karoliina Wehkalampi,[5] Stephen O'Rahilly,[6] Catherine E Aiken,[7] Anthony P Coll,[6] Marcella Ma,[6] Debra Rimmington,[6] Giles SH Yeo,[6] Leo Dunkel[1]*

Affiliations –

[1]Centre for Endocrinology, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK

[2]Centre for Translational Bioinformatics, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK

[3]NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, EC1M 6BQ, UK

[4]Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London, SW7 2AZ, UK

[5]Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, FIN-00029 HUS, Finland

[6]University of Cambridge Metabolic Research Labs and MRC Metabolic Diseases Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK

[7] Department of Obstetrics and Gynaecology, University of Cambridge; NIHR Cambridge Comprehensive Biomedical Research Centre, Cambridge, UK

*Correspondence: l.dunkel@qmul.ac.uk

Word count: 3684

Keywords: Puberty; delayed puberty; FTO; paediatric endocrinology, constitutional delay

Short title: Role of GWAS genes in self-limited delayed puberty

Declaration of interest:

The authors declare no conflict of interest

Funding:

SRH is funded by The Wellcome Trust (102745), Rosetrees Trust (M222) and the Barts and the London Charity (417/1551). AP is funded by Bart's & The London Charity Strategic Research Grant (MEAG1l6R). LG is funded by the Biotechnology and Biological Sciences Research Council (BB/L002671/1). LD is partly supported by the Academy of Finland (14135). MRB and CPC are funded by the National Institutes for Health Research (NIHR), and this work forms part of the portfolio of translational research of the NIHR Biomedical Research Unit at Barts. AD is funded by the MRC (MR/K021613/1). DR, GSHY, SOR and APC are funded by the Medical Research Council (MRC) Metabolic Disease Unit (MRC_MC_UU_12012/1) and animal work was carried out with the assistance of MRC Disease Model Core of the Wellcome Trust MRC Institute of Metabolic Sciences (MRC_MC_UU_12012/5.)

1. I, the designated corresponding author, on behalf of myself and my co-authors, hereby transfer and assign all right, title, and interest, including copyright and any moral rights, in and to the manuscript named in this submission (called the Work hereafter) to the Endocrine Society (ES). If ES ultimately declines to publish the Work in an ES journal, all rights in and to the Work will revert to the author(s).

2. I, and all co-authors, warrant that the Work intended for publication is original and has not been published other than as an abstract or preprint in any language or format and has not been submitted elsewhere for print or electronic publication consideration. We further warrant that the Work does not contain any material that is defamatory or the publication of which would violate any copyright or other personal, intellectual, property, contract, or proprietary right of any person or entity.

3. I warrant that each person listed as an author participated in the Work in a substantive way and is prepared to take public responsibility for it. All authors consent to the investigation of any improprieties that may be alleged regarding the Work. Each author further releases and holds harmless the Endocrine Society from any claim or liability that may arise therefrom.

4. I warrant that I am authorized to accept the terms of this agreement on behalf of myself and all co-authors.

___ I accept the terms of the agreement on behalf of myself and all co-authors.

I certify that neither I nor my co-authors have a conflict of interest as described above that is relevant to the subject matter or materials included in this Work.

Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Abstract

Context: Self-limited delayed puberty (DP) is often associated with delay in physical maturation, but whilst highly heritable the causal genetic factors remain elusive. Genome-wide association studies of the timing of puberty have identified multiple loci for age of menarche (AAM) in females and voice break in males, particularly in pathways controlling energy balance.

Objective/Main outcome measures: We aimed to assess the contribution of rare variants in such genes to the phenotype of familial DP.

Design/Patients: We performed whole exome sequencing (WES) in 67 pedigrees (125 individuals with DP and 35 unaffected controls) from our unique cohort of familial self-limited DP. Using a WES filtering pipeline one candidate gene (*FTO*) was identified. *In silico*, *in vitro* and mouse model studies were performed to investigate the pathogenicity of *FTO* variants and timing of puberty in $FTO^{+/-}$ mice.

Results: We identified potentially pathogenic, rare variants in genes in linkage disequilibrium with GWAS of AAM loci in 283 genes. Of these, 5 genes were implicated in the control of body mass. After filtering for segregation with trait one candidate, *FTO*, was retained. Two *FTO* variants, found in 14 affected individuals from 3 families, were also associated with leanness in these DP patients. One variant (p.Leu44Val) demonstrated altered demethylation activity of the mutant protein *in vitro*. $Fto^{+/-}$ mice displayed a significantly delayed timing of pubertal onset ($p < 0.05$).

Conclusions: Mutations in genes implicated in body mass and timing of puberty in the general population may contribute to the pathogenesis of self-limited DP.

Introduction

Puberty is the maturational process of the reproductive endocrine system that results in adult height and body proportion, in addition to the capacity to reproduce. A minimum level of energy availability is required for the onset of puberty, whilst increased fat mass has been shown to be associated with precocious onset of puberty[1,2]. However, a role for genes connected with regulation of body mass have not been clearly demonstrated in pubertal timing.

The existence of genetic heterogeneity in pubertal timing is supported by several large genome wide association studies (GWAS) of the age of menarche (AAM)[3-5]. Evidence ($P< 5 \times 10^{-8}$) for 123 signals at 106 genomic loci has been identified. Many of these loci were associated with Tanner staging in both sexes, suggesting this data is applicable to both men and women(6,7).

The first of many GWAS loci associated with AAM was the developmental gene *LIN28B*[3,8]. Additional signals in genes involved in energy homeostasis and growth have been found near *LEPR-LEPROT*, which encodes the leptin receptor. Leptin (a key regulator of body mass) is an important permissive signal for the onset of puberty[9]. In addition to leptin signaling, overlap with several genes implicated in body mass index was found, including *FTO*, *SEC16B*, *TMEM18*, and *NEGR1* (Supplementary Table 1) (5). Whether such

genes may regulate pubertal timing exclusively via impact on fat mass or via other BMI-independent mechanisms is unknown [10].

Disordered pubertal timing affects up to 5% of adolescents and is associated with adverse health and psychosocial outcomes [11-14]. Self-limited delayed puberty (DP) represents the extreme end of normal pubertal timing, and is defined as the absence of testicular enlargement in boys or breast development in girls at an age that is 2 to 2.5 standard deviations (SD) later than the population mean [3]. DP may be an isolated feature of the condition or be associated with constitutional delay in growth that can manifest from early childhood.

DP segregates within families, usually with an autosomal dominant pattern of inheritance [15,16]. Despite strong heritability in most cases the genetic basis of DP remains elusive (17) [18]. Moreover, the relevance of genetic factors influencing timing of puberty in the general population to patients with extreme pubertal delay has not been explored. Given the importance of energy balance for reproductive health, genes identified by AAM GWAS that relate to energy homeostasis are of particular interest. Our multi-generational DP families provide a highly valuable resource to investigate these candidate genes in familial DP.

Materials and Methods

Patients

The patients selected for this study are taken from a previously described, accurately phenotyped and characterized, Finnish DP patient cohort [19]. Diagnosis is based on objective evidence of a delayed pubertal growth spurt

148 rather than self-recall. Patients referred with DP to specialist paediatric care in
149 central and southern Finland (1982-2004) were identified. All patients (n=492)
150 met the diagnostic criteria for self-limited DP, defined as the onset of Tanner
151 genital stage II (testicular volume >3 ml) >13.5yr in boys or Tanner breast
152 stage II >13.0yr in girls (i.e. two SD later than average pubertal development)
153 [18,20]. Pubertal growth spurt in probands was more than 2 SD later than
154 average: age at acceleration of pubertal growth (take-off) beyond 13.8 and
155 12.2 yr and age at peak height velocity (PHV) later than 15.6 and 13.7 yr in
156 males and females, respectively (21).
157 Chronic illness and undernutrition was excluded by medical history, clinical
158 examination, and routine laboratory tests. HH, if suspected, was excluded by
159 spontaneous pubertal development at follow-up. In the 50% of patients who
160 choose to have pubertal induction via the use of exogenous sex steroids, all
161 patients were followed up until the point of full pubertal development (Tanner
162 stage G4+ or B4+) to ensure development did not arrest off treatment.
163 Families of the DP patients were invited to participate, with information about
164 medical history and pubertal timing obtained by structured interviews and from
165 archived height records. The criteria for DP in probands' family members were
166 one or more of: 1) age at takeoff or 2) PHV occurring 1.5 SD beyond the
167 mean, i.e. age at takeoff exceeding 12.9 and 11.3 yr, or age
168 at PHV exceeding 14.8 and 12.8 yr in males and females, or 3) age at
169 attaining adult height more than 18 or 16 yr, in males and females,
170 respectively [19]. Previous linkage analysis from this cohort did not find
171 evidence for linked families sharing chromosomal segments identical by

descent, suggesting a founder effect is unlikely to be responsible for this phenotype (19).

Written informed consent was obtained from all participants. The study protocol was approved by the Ethics Committee for Pediatrics, Adolescent Medicine and Psychiatry, Hospital District of Helsinki and Uusimaa (extended to encompass Kuopio, Tampere and Turku University Hospitals) (570/E7/2003). UK ethical approval was granted by the London-Chelsea NRES committee (13/LO/0257). The study was conducted in accordance with the guidelines of The Declaration of Helsinki.

Genetic Analysis

Genetic analysis was performed in 160 individuals from the 67 most extensive families from our cohort with DP. These included 67 probands (male n=57, female n=10), 58 affected family members (male n=36, female n=22) and 35 unaffected family members (male, n=13, female n=22). Whole exome sequencing (WES) was performed on DNA extracted from peripheral blood leukocytes. Variants were analyzed and filtered for potential causal variants in Ingenuity Variant Analysis (Qiagen) using filters for quality control, predicted functional annotation, minor allele frequency (MAF), and GWAS relevance (Figure 1). GWAS relevance filtering allowed identification of those remaining variants that lay within genes in linkage disequilibrium with 106 GWAS loci associated with AAM (n=760) (5). Filters for genes implicated in body mass regulation were applied using a biological context filter with pathway analysis. Variants were filtered for segregation with trait in family members using conventional Sanger sequencing.

196 Targeted exome sequencing using a Fluidigm array of the remaining
197 candidate gene identified post-filtering was then performed in a further 42
198 cohort families (288 individuals, 178 with DP; male=106, female=69 and 110
199 controls; male=55, female=58, Figure 1). Whole gene rare variant burden
200 testing was performed post sequencing.
201 Growth Pattern Analysis
202 The pattern of prepubertal growth in the individuals carrying *FTO* variants was
203 analyzed by using five screening parameters: 1) height for age standard
204 deviation score (HSDS); 2) body mass index (BMI; calculated as weight in
205 kilograms divided by height in meters squared) for age SDS (BMI SDS); 3)
206 HSDS distance from target height (TH) (TH formula = 0.791 × mean parental
207 height SDS − 0.147 for girls and 0.886 × mean parental height SDS – 0.071 for
208 boys; 4) change in height SDS (ΔHSDS); 5) change in BMI SDS (ΔBMI SDS)
209 across time with free age intervals. The calculations of the age-specific and
210 sex-specific normal values for ΔHSDS and ΔBMI SDS were based on
211 longitudinal reference measurements (22). Normality of linear growth was
212 tested by using auxological screening rules based on data from >70,000
213 healthy Finnish children[23].
214 *In silico* Analysis
215 The FTO experimentally solved structure (PDB identifier: 4cxx) was used to
216 study the structural effect of FTO variants. The following interactions involved
217 in protein stability were considered: i) salt bridges; ii) hydrogen bonds (H-
218 bond); and iii) disulphide bridge (S-S bridge). N-glycosylation sites were
219 determined based on the consensus sequence Asn-X-Thr/Ser (X= any amino

acid, except proline). The DSSP program was used to calculate surface accessibility and Disopred3[24] to predict disordered protein regions.

Functional Annotation of FTO mutant proteins

Cloning of wild-type human FTO cDNA into pET302/NT-His has been described previously[25]. The p.Leu44Val and p.Ala163Thr point mutations were introduced using PCR-mediated mutagenesis (Quickchange II, Agilent Technoligies) using primers FTO_L44V FOR: 5'-GAATTCTATCAGCAGTGGCAGGTGAAATATCCTAAACTAATTCT-3', REV: 5'-AGAATTAGTTTAGGATATTTCACCTGCCACTGCTGATAGAATTC-3' and FTO_A163T FOR: 5'-CACAGCATCCTCATTAGTCTTCTCTTTGGCAGCAA-3', REV: 5'-TTGCTGCCAAAGAGAAGACTAATGAGGATGCTGTG-3' and verified by sequencing. An RNase-cleavage assay[26] was used to measure the demethylation activity of FTO on 3-methyl-uridine (3-meU). Recombinant wild-type and mutant FTO expression plasmids were transformed into *Escherichia coli* BL21-Gold (DE3) (Stratagene) and cultured in LB broth and 50 μg/ml carbenicillin. Expression of the cloned gene was induced by the addition of IPTG (isopropyl-β-D-1-thiogalactopyranoside) at 1 mM final concentration at 15◦C for 4 h. The cells were harvested and pellets resuspended in lysis buffer [50 mM HEPES-KOH (pH 8.0), 2 mM 2-mercaptoethanol, 5% glycerol and 300mM NaCl] before digestion with lysozyme (1 mg/ml). The cleared lysate was supplemented with imidazole (final concentration 10 mM) before mixing with 1 ml of pre- washed Ni-NTA (Ni2+-nitrilotriacetate) beads (Qiagen). After binding for 1 h in the cold, the mixture was washed with lysis buffer supplemented with increasing concentrations of imidazole. FTO was eluted with 2 ml of lysis buffer

245 containing 250 mM imidazole. The eluate was concentrated with a 30 kDa
246 molecular-mass cut-off concentrator (Sartorius Stedim) with buffer changing
247 to 20 mM HEPES-KOH (pH 8), 5 % glycerol and 50 mM NaCl. Purified
248 proteins were snap-frozen and stored at −80◦C. Protein purity was estimated
249 by Commassie Blue staining after resolving by SDS/PAGE (4–12 % gradient
250 gels; Invitrogen).
251 Dose response of FTO on 3-meU demethylation: Recombinant FTO proteins
252 were assayed as previously described[26]. Each protein, at different protein
253 concentrations from 0 -1000 nM, was assayed in a reaction containing 100
254 nM substrate, 75 μM Fe(NH4)2(SO4)2, 300 μM 2-OG, 2 mM ascorbate,
255 50 μg/ml BSA and 62.5 μg/μl of RNase A in 50 mM Tris/HCl buffer at pH 7.0.
256 Samples were prepared in duplicate in a dark flat-bottomed 96-well plate and
257 the FAM (6-carboxyfluorescein) emission was measured for 30 min at a
258 wavelength of 520 nm with excitation at 485 nm. The measurement was
259 performed at room temperature (25°C) using a microplate reader [Infinite
260 M1000, Tecan]. Wild type (WT) FTO protein and catalytically inactive mutant
261 p.Arg316Gln (R316Q) served as positive and negative controls respectively.
262 Mouse experiments
263 Fto deficient mice were a generous gift from Prof. Roger Cox (MRC Harwell,
264 Oxford) and were genotyped as previously described (27). This research is
265 regulated under the Animals (Scientific Procedures) Act 1986 Amendment
266 Regulations 2012 following ethical review by the University of Cambridge
267 Animal Welfare and Ethical Review Body (AWERB). Animals were kept under
268 controlled temperature (22°C) and a 12-h light, 12-h dark schedule (lights on

07:00–19:00). Standard chow (Special Diet Services) and water were available *ad libitum*.

For the vaginal opening study female *Fto* heterozygous mice ($Fto^{+/-}$) (n=45) and their WT littermates (n=24) were taken from either a male *Fto* WT x female $Fto^{+/-}$ cross or a male $Fto^{+/-}$ x female *Fto* WT cross. From P21 (day of weaning) all female mice were weighed and visual examination of the vagina was carried out by placing the mouse on top of a cage lid and lifting the tail vertically away from the body. No excessive force was involved. First day of vaginal opening was recorded when a complete opening was observed.

For all experiments, data are expressed as the mean ± SEM. To determine statistical significance, we used the unpaired t test (2-tailed) using SPSS Software (version 24). A p value of <0.05 was considered statistically significant.

Results

Variants in GWAS genes implicated in body mass were identified following exome sequencing in families with self-limited delayed puberty

WES performed in the 67 largest and best phenotyped families from our cohort (160 individuals: a total of 125 individuals with DP, male=93, female=32; and 35 controls, male=13, female=22), identified 6,952,773 variants after quality control (Figure 1). Filtering to identify high quality, rare, predicted deleterious variants not present in control subjects selected 12,371 variants in 7,470 genes. Of these 7,470 genes, 238 were found to be in linkage disequilibrium with a GWAS locus for timing of puberty, and 5 of these

238 were genes implicated in body mass regulation or growth by pathway analysis. Of these 5 genes, 4 (*GPD2*, *GHR*, *ESR1* and *VDR*) were found to have only variants that did not segregate with the DP trait in family members. The remaining candidate gene, *FTO* (Fat mass and obesity-associated protein, ENSG00000140718, gene identification number 79068), has been previously described in the literature as involved in pathways of energy homeostasis and growth[5], and is known to act as an Fe(II) 2-OG (2-oxoglutarate) -dependent dioxygenase to repair alkylated DNA and RNA by demethylation[26]. *FTO* contributes to the regulation of energy balance, and thus to the regulation of body size and fat accumulation.

Two variants in *FTO* (NM_001080432.2: c.130C>G p.Leu44Val and NM_001080432.2: c.487G>A (rs145884431) p.Ala163Thr) were identified in three families from our cohort and found in one or fewer control subjects (rare variant burden testing adjusted p = 0.058). Both variants are rare (MAF < 0.2%) heterozygous missense variants and predicted benign or tolerated by >2/5 prediction software tools.

Families with potentially pathogenic *FTO* variants display autosomal dominant inheritance of DP phenotype and low body mass.

The family identified with the p.Ala163Thr variant (family 1) and both of the families with the p.Leu44Val variant (families 2 and 3) displayed the typical autosomal inheritance pattern of the DP trait, with perfect segregation (Figure 2, panel A). Affected individuals from family 1 with the p.Ala163Thr variant and from family 3 with the p.Leu44Val variant were particularly underweight in childhood, with the two probands from these families (individuals 1.III.2 and 3.III.2) falling into the thinness grade 2 category(28) before puberty (Figure 2,

panels B and D). Although there was some variability in this phenotype, all family members carrying *FTO* variants had ISO-BMI values in the lower range (<23) (Figure 2 and Supplementary Fig. 1-3, Table 1). In addition, both of the probands from families 2 and 3 who carry the p.Leu44Val displayed faltering growth in early childhood. Both displayed significant deflection from previous height measurements in the 2 years following birth, as well as height significantly below target height in later adolescence associated with delayed pubertal growth (Figure 2, panels C and D)[22].

In silico analysis of potential mutations

We carried out *in silico* analysis using the solved structure of FTO (PDB identifier: 3lfm) to determine the possible pathogenicity of the identified variants. The hydrophobic residue Leucine 44 is part of a solvent-exposed alpha helix on the surface. Substitution with Valine is not predicted to alter the structure of FTO or interaction with iron molecules or DNA. However, L44 and other residues in the same solvent-exposed alpha helix form a motif (Supplementary Fig. 4 and 5), which is highly conserved across placental mammals but not reptiles, birds or fish (Supplementary Fig. 6). This motif (residues 36-48) forms a patch on the FTO protein surface (Supplementary Fig. 7). This may act as a mammal-specific interaction site (between FTO and another protein), required for FTO function for example in reproductive development. In this scenario, a small change in side chain volume, such as Leucine-to-Valine, may have a subtle effect in protein-protein interaction and lead to a change in FTO activity *in vivo*.

Alanine 163 is a hydrophobic, not highly conserved residue (Supplementary Fig. 8, panel A). Alanine 163 is at the end of the H4 alpha helix and the

beginning of a long, disordered region (Supplementary Fig. 8, panel B), which connects helices H4 and H5 (Supplementary Fig. 8, panel C).

FTO p.Leu44Val mutant protein displays reduced demethylase activity *in vitro*

We carried out functional characterization of the identified mutant FTO proteins (p.Leu44Val and p.Ala163Thr) as compared to WT protein. A previously verified RNase-cleavage assay was used to measure the demethylation activity of FTO on 3-meU (26). Although kinetic activity of the mutant protein p.Ala163Thr did not vary from WT using this assay, mutant protein p.Leu44Val showed an approximately 20% lower kinetic activity than WT activity (Figure 3).

***FTO* deficiency *in vivo* results in delayed vaginal opening in mice**

In order to examine the influence of FTO activity on pubertal timing in an *in vivo* model, we examined timing of puberty in mice deficient for FTO in the heterozygous state (*Fto*$^{+/-}$), in keeping with the human genotype identified. *Fto*$^{-/-}$ mice were not selected for these experiments because of their poor postnatal health (29). *Fto*$^{+/-}$ mice had significantly delayed timing of vaginal opening (VO) (mean postnatal day +/- SEM: 27.20 +/- 0.44 in wild-type (n=24) vs 28.56 +/- 0.48 in *Fto*$^{+/}$ mice (n=45), p =0.047), an event which reflects the pubertal rise in estradiol [30] (Figure 4). Mean body weight of the *Fto*$^{+/}$ group was not significantly different to the WT mice (mean body weight (in g) +/- SEM: 11.64 +/- 0.21 in wild-type vs 11.45 +/- 0.14 in *Fto*$^{+/}$ mice, p=0.467) (Figure 5).

Using simple linear modelling, *Fto* genotype of the pup (Het vs WT) explained approximately 3% of the total variation in timing of VO. Consideration of an

369 additional factor, maternal genotype, improved the model by increasing the
370 significance of the association between pup genotype and timing of VO
371 slightly (p=0.04), and accounted for 6% of the total variation in timing of VO.
372 In contrast, paternal genotype decreased the significance and total variation
373 accounted for by the model.
374
375 **Discussion**
376 Genome wide association studies of AAM in the general population have
377 attempted to unravel the complex conundrum of which genetic
378 factors influence the timing of puberty. Despite many loci being identified,
379 clear evidence for the role of particular genes and pathways is for the most
380 part lacking. Those genes lying within pathways of energy metabolism and
381 growth appear promising, with the discovery of the role of *Lin28B* in C.elegans
382 development[3] and the importance of leptin as a permissive signal in triggering
383 the onset of puberty(9,31).
384 The inheritance of DP is known to be under strong genetic influence with
385 commonly an autosomal dominant inheritance pattern, and thus represents a
386 useful basis for the investigation of puberty genetics. Notably, self-limited or
387 constitutional DP is often associated with slow maturation throughout
388 childhood, implicating growth and energy metabolism pathways in its
389 pathogenesis. Previously, genes in such pathways identified through GWAS
390 have not been screened in patients with DP.
391 Our results have identified variants in *FTO* as a potential contributory factor in
392 the development of self-limited DP in three pedigrees from our large cohort of
393 patients with familial DP. *FTO* (fat mass and obesity associated gene) was the

394 first obesity-susceptibility gene identified through GWAS and continues to be
395 the locus with the largest effect on body mass index (BMI) and obesity risk[10].
396 Those DP patients identified with *FTO* variants from our study showed
397 reductions in body mass. The *FTO* variants carried by our DP patients may
398 result in reduced fat mass, which would in turn contribute to a delay in the
399 timing of pubertal onset. This delay may be mediated directly through reduced
400 leptin levels. Although we do not routinely measure leptin levels in DP
401 patients, leptin levels have been shown to be significantly lower in pubertal-
402 age patients with self-limited DP(32).
403 Notably, in an *in vivo* model $Fto^{+/-}$ mice had a significantly delayed onset of
404 puberty as compared to WT mice. In the 7 days preceding puberty onset,
405 however, body weight was not significantly different between the two pup
406 genotype groups. Previous studies have demonstrated that $Fto^{-/-}$ mice show a
407 30-40% reduction in body weight by 6 weeks of age[29] and that transgenic
408 mice with additional copies of *Fto* show a dose-dependent increase in body
409 and fat mass[33]. However, the relationship between *FTO* genotype, fat mass
410 and leptin levels remains somewhat unclear. *Fto* deficient mice do become
411 obese when subjected to a high fat diet, although they remain sensitive to the
412 anorexigenic effects of leptin (29,34).
413 Moreover, it is possible that *FTO* gene dosage may have an effect on energy
414 homeostasis independent of effects on fat mass[33], including on the balance
415 between catabolic and anabolic pathways (35). *FTO* has been identified as an
416 amino acid sensor acting, via mTOR, to influence appropriate levels of
417 development and translation[36]. *FTO* is expressed within the hypothalamus in
418 several sites critical for energy balance, including in the arcuate nucleus

419 within proopiomelanocortin (POMC) neurons(37,38). In one study *Fto* levels in
420 the arcuate nuclei of fasted mice fell by up to 60%, and this was not rescued
421 by leptin administration. Other studies have shown conflicting results in the
422 effects on *Fto* mRNA levels of fasting, depending on whether whole
423 hypothalamus or arcuate nucleus were studied and on the length of fast (38).
424 However, *Fto*$^{-/-}$ mice display blunted starvation-induced Npy mRNA
425 induction[29]. More recent studies have suggested that Fto may influence the
426 metabolic outcomes of a high fat diet via hypothalamic signaling pathways
427 acting independently of body weight (34). Mutations in *FTO*, including those
428 with greatly reduced demethylase activity (e.g. pR316Q, Figure 3), have been
429 identified in human subjects associated with both lean and obese phenotypes
430 [25]. We were not able in our study to identify the mechanism by which the
431 p.Ala163Thr variant might affect protein function; although no reduction in
432 demethylation activity was demonstrated it is possible that this variant may
433 produce a deleterious effect by another route, for example defects in post-
434 translational modification or protein degradation.
435 Thus, FTO may be important for signaling energy sufficiency
436 and the 'healthy energy balance' required for pubertal onset. Our *in*
437 *silico* analysis suggests that the p.Leu44Val mutation we have identified may
438 represent a mammal-specific interaction site between FTO and another
439 protein (or DNA), important for FTO function in terms of reproductive
440 development. Moreover, maternal genotype may contribute to pubertal timing,
441 as demonstrated from our *Fto*$^{+/-}$ mice data. A reproductive phenotype present
442 in *Fto* heterozygote mothers could expose pups to a suboptimal environment
443 that could influence their puberty timing.

Finally, our finding of maturational delay in growth in early childhood in the two probands with p.Leu44Val mutation is of interest. Constitutional delay in growth is seen in a subset of patients with DP, and our findings implicate mutations in energy pathway genes in the pathogenesis of patients with such a phenotype.

Overall, our discovery of two rare variants in *FTO* associated with self-limited DP in our large familial cohort, and of delayed vaginal opening in *FTO*-deficient mice, provides evidence that perturbations in pathways of energy homeostasis and growth may potentially produce a phenotype of DP. We note that despite this extensive analysis, only three of 67 probands were identified with potentially pathogenic variants in such pathways, highlighting the high degree of heterogeneity in the genetic basis of self-limited DP. These findings merit further exploration in our own cohort and in other populations, including sub-group analysis of DP patients with low BMI from early childhood.

Acknowledgements

We are very grateful to the patients and families who contributed their time, medical information, and DNA samples to this study.

References

1. Kaplowitz PB, Slora EJ, Wasserman RC, Pedlow SE, Herman-Giddens ME. Earlier onset of puberty in girls: relation to increased body mass index and race. Pediatrics 2001; 108:347-353

2. He Q, Karlberg J. Bmi in childhood and its association with height gain, timing of puberty, and final height. Pediatric research 2001; 49:244-251

3. Ong KK, Elks CE, Li S, Zhao JH, Luan J, Andersen LB, Bingham SA, Brage S, Smith GD, Ekelund U, Gillson CJ, Glaser B, Golding J, Hardy R, Khaw KT, Kuh D, Luben R, Marcus M, McGeehin MA, Ness AR, Northstone K, Ring SM, Rubin C, Sims MA, Song K, Strachan DP, Vollenweider P, Waeber G, Waterworth DM, Wong A, Deloukas P, Barroso I, Mooser V, Loos RJ, Wareham NJ. Genetic variation in LIN28B is associated with the timing of puberty. Nature genetics 2009; 41:729-733

4. Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, Lunetta KL, Visser JA, Byrne EM, Cousminer DL, Gudbjartsson DF, Esko T, Feenstra B, Hottenga JJ, Koller DL, Kutalik Z, Lin P, Mangino M, Marongiu M, McArdle PF, Smith AV, Stolk L, van Wingerden SH, Zhao JH, Albrecht E, Corre T, Ingelsson E, Hayward C, Magnusson PK, Smith EN, Ulivi S, Warrington NM, Zgaga L, Alavere H, Amin N, Aspelund T, Bandinelli S, Barroso I, Berenson GS, Bergmann S, Blackburn H, Boerwinkle E, Buring JE, Busonero F, Campbell H, Chanock SJ, Chen W, Cornelis MC, Couper D, Coviello AD, d'Adamo P, de Faire U, de Geus EJ, Deloukas P, Doring A, Smith GD, Easton DF, Eiriksdottir G, Emilsson V, Eriksson J, Ferrucci L, Folsom AR, Foroud T, Garcia M, Gasparini P, Geller F, Gieger C, Gudnason V, Hall P, Hankinson SE, Ferreli L, Heath AC, Hernandez DG, Hofman A, Hu FB, Illig T, Jarvelin MR, Johnson AD, Karasik D, Khaw KT, Kiel DP, Kilpelainen TO, Kolcic I, Kraft P, Launer LJ, Laven JS, Li S, Liu J, Levy D, Martin NG, McArdle WL, Melbye M, Mooser V, Murray JC, Murray SS, Nalls MA, Navarro P, Nelis M, Ness AR, Northstone K, Oostra BA, Peacock M, Palmer LJ, Palotie A, Pare G, Parker AN, Pedersen NL, Peltonen L, Pennell CE, Pharoah P, Polasek O, Plump AS, Pouta A, Porcu E, Rafnar T, Rice JP, Ring SM, Rivadeneira F, Rudan I, Sala C, Salomaa V, Sanna S, Schlessinger D, Schork NJ, Scuteri A, Segre AV, Shuldiner AR, Soranzo N, Sovio U, Srinivasan SR, Strachan DP, Tammesoo ML, Tikkanen E, Toniolo D, Tsui K, Tryggvadottir L, Tyrer J, Uda M, van Dam RM, van Meurs JB, Vollenweider P, Waeber G, Wareham NJ, Waterworth DM, Weedon MN, Wichmann HE, Willemsen G, Wilson JF, Wright AF, Young L, Zhai G, Zhuang WV, Bierut LJ, Boomsma DI, Boyd HA, Crisponi L, Demerath EW, van Duijn CM, Econs MJ, Harris TB, Hunter DJ, Loos RJ, Metspalu A, Montgomery GW, Ridker PM, Spector TD, Streeten EA, Stefansson K, Thorsteinsdottir U, Uitterlinden AG, Widen E, Murabito JM, Ong KK, Murray A. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nature genetics 2010; 42:1077-1085

5. Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E, Ang WQ, Corre T, Cousminer DL, Feenstra B, Franceschini N, Ganna A, Johnson AD, Kjellqvist S, Lunetta KL,

McMahon G, Nolte IM, Paternoster L, Porcu E, Smith AV, Stolk L, Teumer A, Tsernikova N, Tikkanen E, Ulivi S, Wagner EK, Amin N, Bierut LJ, Byrne EM, Hottenga JJ, Koller DL, Mangino M, Pers TH, Yerges-Armstrong LM, Hua Zhao J, Andrulis IL, Anton-Culver H, Atsma F, Bandinelli S, Beckmann MW, Benitez J, Blomqvist C, Bojesen SE, Bolla MK, Bonanni B, Brauch H, Brenner H, Buring JE, Chang-Claude J, Chanock S, Chen J, Chenevix-Trench G, Collee JM, Couch FJ, Couper D, Coviello AD, Cox A, Czene K, D'Adamo A P, Davey Smith G, De Vivo I, Demerath EW, Dennis J, Devilee P, Dieffenbach AK, Dunning AM, Eiriksdottir G, Eriksson JG, Fasching PA, Ferrucci L, Flesch-Janys D, Flyger H, Foroud T, Franke L, Garcia ME, Garcia-Closas M, Geller F, de Geus EE, Giles GG, Gudbjartsson DF, Gudnason V, Guenel P, Guo S, Hall P, Hamann U, Haring R, Hartman CA, Heath AC, Hofman A, Hooning MJ, Hopper JL, Hu FB, Hunter DJ, Karasik D, Kiel DP, Knight JA, Kosma VM, Kutalik Z, Lai S, Lambrechts D, Lindblom A, Magi R, Magnusson PK, Mannermaa A, Martin NG, Masson G, McArdle PF, McArdle WL, Melbye M, Michailidou K, Mihailov E, Milani L, Milne RL, Nevanlinna H, Neven P, Nohr EA, Oldehinkel AJ, Oostra BA, Palotie A, Peacock M, Pedersen NL, Peterlongo P, Peto J, Pharoah PD, Postma DS, Pouta A, Pylkas K, Radice P, Ring S, Rivadeneira F, Robino A, Rose LM, Rudolph A, Salomaa V, Sanna S, Schlessinger D, Schmidt MK, Southey MC, Sovio U, Stampfer MJ, Stockl D, Storniolo AM, Timpson NJ, Tyrer J, Visser JA, Vollenweider P, Volzke H, Waeber G, Waldenberger M, Wallaschofski H, Wang Q, Willemsen G, Winqvist R, Wolffenbuttel BH, Wright MJ, Australian Ovarian Cancer S, Network G, kConFab, LifeLines Cohort S, InterAct C, Early Growth Genetics C, Boomsma DI, Econs MJ, Khaw KT, Loos RJ, McCarthy MI, Montgomery GW, Rice JP, Streeten EA, Thorsteinsdottir U, van Duijn CM, Alizadeh BZ, Bergmann S, Boerwinkle E, Boyd HA, Crisponi L, Gasparini P, Gieger C, Harris TB, Ingelsson E, Jarvelin MR, Kraft P, Lawlor D, Metspalu A, Pennell CE, Ridker PM, Snieder H, Sorensen TI, Spector TD, Strachan DP, Uitterlinden AG, Wareham NJ, Widen E, Zygmunt M, Murray A, Easton DF, Stefansson K, Murabito JM, Ong KK. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 2014; 514:92-97

6. Day FR, Bulik-Sullivan B, Hinds DA, Finucane HK, Murabito JM, Tung JY, Ong KK, Perry JR. Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. Nat Commun 2015; 6:8842

7. Cousminer DL, Stergiakouli E, Berry DJ, Ang W, Groen-Blokhuis MM, Korner A, Siitonen N, Ntalla I, Marinelli M, Perry JR, Kettunen J, Jansen R, Surakka I, Timpson NJ, Ring S, McMahon G, Power C, Wang C, Kahonen M, Viikari J, Lehtimaki T, Middeldorp CM, Hulshoff Pol HE, Neef M, Weise S, Pahkala K, Niinikoski H, Zeggini E, Panoutsopoulou K, Bustamante M, Penninx BW, ReproGen C, Murabito J, Torrent M, Dedoussis GV, Kiess W, Boomsma DI, Pennell CE, Raitakari OT, Hypponen E, Davey Smith G, Ripatti S, McCarthy MI, Widen E, Early Growth Genetics C. Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. Human molecular genetics 2014; 23:4452-4464

8. Perry JR, Stolk L, Franceschini N, Lunetta KL, Zhai G, McArdle PF, Smith AV, Aspelund T, Bandinelli S, Boerwinkle E, Cherkas L, Eiriksdottir G,

559 Estrada K, Ferrucci L, Folsom AR, Garcia M, Gudnason V, Hofman A,
560 Karasik D, Kiel DP, Launer LJ, van Meurs J, Nalls MA, Rivadeneira F,
561 Shuldiner AR, Singleton A, Soranzo N, Tanaka T, Visser JA, Weedon MN,
562 Wilson SG, Zhuang V, Streeten EA, Harris TB, Murray A, Spector TD,
563 Demerath EW, Uitterlinden AG, Murabito JM. Meta-analysis of genome-
564 wide association data identifies two loci influencing age at menarche.
565 Nature genetics 2009; 41:648-650
566 **9.** Barash IA, Cheung CC, Weigle DS, Ren H, Kabigting EB, Kuijper JL, Clifton
567 DK, Steiner RA. Leptin is a metabolic signal to the reproductive system.
568 Endocrinology 1996; 137:3144-3147
569 **10.** Yeo GS. The role of the FTO (Fat Mass and Obesity Related) locus in
570 regulating body size and composition. Molecular and cellular
571 endocrinology 2014; 397:34-41
572 **11.** Widen E, Silventoinen K, Sovio U, Ripatti S, Cousminer DL, Hartikainen AL,
573 Laitinen J, Pouta A, Kaprio J, Jarvelin MR, Peltonen L, Palotie A. Pubertal
574 timing and growth influences cardiometabolic risk factors in adult males
575 and females. Diabetes care 2012; 35:850-856
576 **12.** Ritte R, Lukanova A, Tjonneland A, Olsen A, Overvad K, Mesrine S,
577 Fagherazzi G, Dossus L, Teucher B, Steindorf K, Boeing H, Aleksandrova K,
578 Trichopoulou A, Lagiou P, Trichopoulos D, Palli D, Grioni S, Mattiello A,
579 Tumino R, Sacerdote C, Quiros JR, Buckland G, Molina-Montes E, Chirlaque
580 MD, Ardanaz E, Amiano P, Bueno-de-Mesquita B, van Duijnhoven F, van
581 Gils CH, Peeters PH, Wareham N, Khaw KT, Key TJ, Travis RC, Krum-
582 Hansen S, Gram IT, Lund E, Sund M, Andersson A, Romieu I, Rinaldi S,
583 McCormack V, Riboli E, Kaaks R. Height, age at menarche and risk of
584 hormone receptor positive and negative breast cancer: A cohort study.
585 International journal of cancer Journal international du cancer 2012;
586 **13.** He C, Zhang C, Hunter DJ, Hankinson SE, Buck Louis GM, Hediger ML, Hu
587 FB. Age at menarche and risk of type 2 diabetes: results from 2 large
588 prospective cohort studies. American journal of epidemiology 2010;
589 171:334-344
590 **14.** Day FR, Elks CE, Murray A, Ong KK, Perry JR. Puberty timing associated
591 with diabetes, cardiovascular disease and also diverse health outcomes in
592 men and women: the UK Biobank study. Sci Rep 2015; 5:11208
593 **15.** Sedlmeyer IL. Pedigree Analysis of Constitutional Delay of Growth and
594 Maturation: Determination of Familial Aggregation and Inheritance
595 Patterns. Journal of Clinical Endocrinology & Metabolism 2002; 87:5581-
596 5586
597 **16.** Wehkalampi K, Widen E, Laine T, Palotie A, Dunkel L. Patterns of
598 inheritance of constitutional delay of growth and puberty in families of
599 adolescent girls and boys referred to specialist pediatric care. The Journal
600 of clinical endocrinology and metabolism 2008; 93:723-728
601 **17.** Gajdos ZK, Hirschhorn JN, Palmert MR. What controls the timing of
602 puberty? An update on progress from genetic investigation. Current
603 opinion in endocrinology, diabetes, and obesity 2009; 16:16-24
604 **18.** Palmert MR, Dunkel L. Clinical practice. Delayed puberty. N Engl J Med
605 2012; 366:443-453

606 **19.** Wehkalampi K, Widen E, Laine T, Palotie A, Dunkel L. Association of the
607 timing of puberty with a chromosome 2 locus. The Journal of clinical
608 endocrinology and metabolism 2008; 93:4833-4839
609 **20.** Sadov S, Koskenniemi JJ, Virtanen HE, Perheentupa A, Petersen JH,
610 Skakkebaek NE, Main KM, Toppari J. Testicular Growth During Puberty in
611 Boys With and Without a History of Congenital Cryptorchidism. The
612 Journal of clinical endocrinology and metabolism 2016; 101:2570-2577
613 **21.** Tanner JM, Whitehouse RH, Marubini E, Resele LF. The adolescent growth
614 spurt of boys and girls of the Harpenden growth study. Ann Hum Biol
615 1976; 3:109-126
616 **22.** Saari A, Harju S, Makitie O, Saha MT, Dunkel L, Sankilampi U. Systematic
617 growth monitoring for the early detection of celiac disease in children.
618 JAMA Pediatr 2015; 169:e1525
619 **23.** Saari A, Sankilampi U, Hannila ML, Kiviniemi V, Kesseli K, Dunkel L. New
620 Finnish growth references for children and adolescents aged 0 to 20
621 years: Length/height-for-age, weight-for-length/height, and body mass
622 index-for-age. Ann Med 2011; 43:235-248
623 **24.** Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions
624 with annotated protein-binding activity. Bioinformatics 2015; 31:857-863
625 **25.** Meyre D, Proulx K, Kawagoe-Takaki H, Vatin V, Gutierrez-Aguilar R, Lyon
626 D, Ma M, Choquet H, Horber F, Van Hul W, Van Gaal L, Balkau B, Visvikis-
627 Siest S, Pattou F, Farooqi IS, Saudek V, O'Rahilly S, Froguel P, Sedgwick B,
628 Yeo GS. Prevalence of loss-of-function FTO mutations in lean and obese
629 individuals. Diabetes 2010; 59:311-318
630 **26.** Ma M, Harding HP, O'Rahilly S, Ron D, Yeo GS. Kinetic analysis of FTO (fat
631 mass and obesity-associated) reveals that it is unlikely to function as a
632 sensor for 2-oxoglutarate. The Biochemical journal 2012; 444:183-187
633 **27.** McMurray F, Church CD, Larder R, Nicholson G, Wells S, Teboul L, Tung
634 YC, Rimmington D, Bosch F, Jimenez V, Yeo GS, O'Rahilly S, Ashcroft FM,
635 Coll AP, Cox RD. Adult onset global loss of the fto gene alters body
636 composition and metabolism in the mouse. PLoS genetics 2013;
637 9:e1003166
638 **28.** Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to
639 define thinness in children and adolescents: international survey. Bmj
640 2007; 335:194
641 **29.** Fischer J, Koch L, Emmerling C, Vierkotten J, Peters T, Bruning JC, Ruther
642 U. Inactivation of the Fto gene protects from obesity. Nature 2009;
643 458:894-898
644 **30.** Nelson JF, Karelus K, Felicio LS, Johnson TE. Genetic influences on the
645 timing of puberty in mice. Biology of reproduction 1990; 42:649-655
646 **31.** Farooqi IS, Jebb SA, Langmack G, Lawrence E, Cheetham CH, Prentice AM,
647 Hughes IA, McCamish MA, O'Rahilly S. Effects of recombinant leptin
648 therapy in a child with congenital leptin deficiency. N Engl J Med 1999;
649 341:879-884
650 **32.** Gill MS, Hall CM, Tillmann V, Clayton PE. Constitutional delay in growth
651 and puberty (CDGP) is associated with hypoleptinaemia. Clinical
652 endocrinology 1999; 50:721-726
653 **33.** Church C, Moir L, McMurray F, Girard C, Banks GT, Teboul L, Wells S,
654 Bruning JC, Nolan PM, Ashcroft FM, Cox RD. Overexpression of Fto leads

655 to increased food intake and results in obesity. Nature genetics 2010;
656 42:1086-1092
657 **34.** Tung YC, Gulati P, Liu CH, Rimmington D, Dennis R, Ma M, Saudek V,
658 O'Rahilly S, Coll AP, Yeo GS. FTO is necessary for the induction of leptin
659 resistance by high-fat feeding. Molecular metabolism 2015; 4:287-298
660 **35.** Merkestein M, McTaggart JS, Lee S, Kramer HB, McMurray F, Lafond M,
661 Boutens L, Cox R, Ashcroft FM. Changes in gene expression associated
662 with FTO overexpression in mice. PloS one 2014; 9:e97162
663 **36.** Speakman JR. The 'Fat Mass and Obesity Related' (FTO) gene:
664 Mechanisms of Impact on Obesity and Energy Balance. Curr Obes Rep
665 2015; 4:73-91
666 **37.** Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, Hewitson KS, Yeo GS,
667 McDonough MA, Cunliffe S, McNeill LA, Galvanovskis J, Rorsman P, Robins
668 P, Prieur X, Coll AP, Ma M, Jovanovic Z, Farooqi IS, Sedgwick B, Barroso I,
669 Lindahl T, Ponting CP, Ashcroft FM, O'Rahilly S, Schofield CJ. The obesity-
670 associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid
671 demethylase. Science 2007; 318:1469-1472
672 **38.** McTaggart JS, Lee S, Iberl M, Church C, Cox RD, Ashcroft FM. FTO is
673 expressed in neurones throughout the brain and its expression is
674 unaltered by fasting. PloS one 2011; 6:e27968
675

Figure Legends:

Figure 1 – Flowchart of WES (whole exome sequencing) filtering strategy to identify candidate genes.

Whole exome sequencing was initially performed on DNA extracted from peripheral blood leukocytes of 160 individuals from the 67 most extensive families from our cohort (125 with DP and 35 controls), with exome capture on a Nimblegen V2 or Agilent V5 platform and sequencing on the Illumina Hiseq 2000. The exome sequences were aligned to the UCSC hg19 reference genome. Picard tools and the genome analysis toolkit were used to mark PCR duplicates, realign around indels, recalibrate quality scores and call variants. Variants were then analyzed further and filtered for potential causal variants using filters for quality control, predicted functional annotation, minor allele frequency (MAF), segregation with trait and GWAS relevance (See methods for further information on filtering criteria). Targeted exome sequencing using a Fluidigm array of a candidate gene identified post-filtering was then performed in a further 42 families from the same cohort (288 individuals, 178 with DP and 110 controls). Variants post targeted re-sequencing were filtered using the same criteria as the whole exome sequencing data. Functional annotation of the variants as described elsewhere in methods. DP – delayed puberty.

Figure 2 – Pedigrees and auxological data of the families with potentially pathogenic *FTO* variants

Panel A: Squares indicate male family members, circles female family members. Black symbols represent clinically affected, grey represent unknown phenotype, clear symbols represent unaffected individuals. The arrow with 'P' indicates the proband in each family and 'us' indicates un-sequenced due to lack of DNA from that individual. The mutation in each family is given next to the family number; a horizontal black line above an individual's symbol indicates they are heterozygous for the variant as confirmed by either whole exome sequencing or Fluidigm array, and verified by Sanger sequencing. A red dot indicates the individual was underweight (thinness grade 2 or more significant) and '?' indicates that BMI information for that individual is not available.

Panels B-D: BMI and height standard deviation score (SDS) charts for the probands of each of the three pedigrees (family 1.III.2, family 2.III.5 and family 3.III.2). Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published [22].

Figure 3 – Demethylation assay assessing kinetic activity of mutant versus wild type FTO proteins.

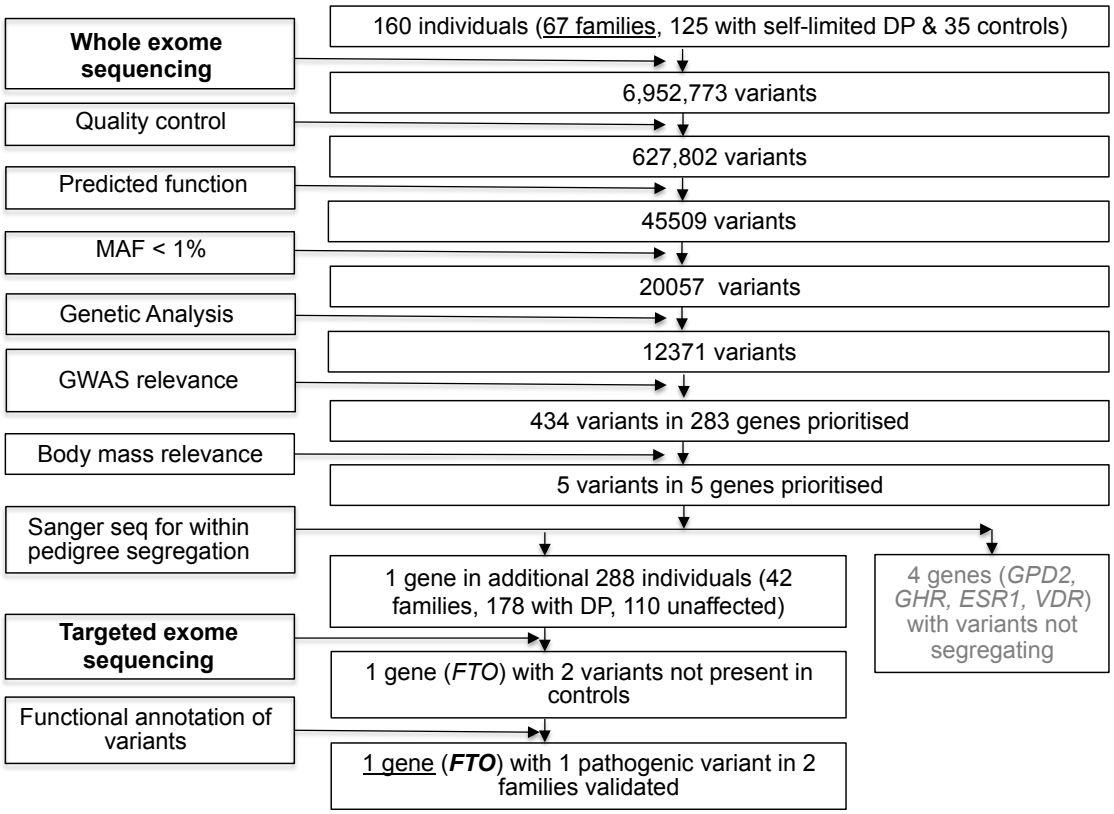
FTO activity is proportional to the concentration present in the reaction. Demethylase activity is likely to be related to the ability of FTO to function as a sensor for cellular metabolism (36). The R316Q mutant is enzymatically dead across all concentrations tested. The A163T and L44V mutants showed

725 demethylase activity towards methylated-uridine in a dose-dependent manner
726 but with different affinities. WT – wild-type
727
728 **Figure 4 – Timing of vaginal opening in wild-type (WT) and *FTO*$^{+/-}$**
729 **heterozygous (Het) mice.**
730 Cumulative percentages of mice displaying vaginal opening by postnatal day
731 are shown for WT and *FTO*$^{+/-}$ mice. WT mice n=24, *FTO*$^{+/-}$ n=45; p <0.05 by
732 un-paired t test.
733
734 **Figure 5 – Mean body weight (g) for wild type (WT) and *Fto*$^{+/-}$ (Het) mice**
735 **in 7 days prior to vaginal opening**
736 Mean body weight (g) +/- SEM: 11.64 +/- 0.21 in wild-type (n=24) vs 11.45 +/-
737 0.14 in *Fto*$^{+/}$ mice (n=45), p=0.467 by un-paired t test. Error bars show SEM
738 for each group each day.
739
740

Case	Sex	Amino acid alteration	Height SDS at age 4 yrs	Height SDS at age 8/9 yrs	Height SDS at age 18 yrs	ISO-BMI at 18 yrs
1.II.1	M	p.Ala163Thr	-	1.1	1.7	16.9
1.III.2 (P)	M	p.Ala163Thr	1.1	0.5	1.1	17.1
1.III.1	F	p.Ala163Thr	0.9	1.0	1.1	17.3
1.II.5	M	p.Ala163Thr	-1.0	-1.0	-0.4	-
2.III.5 (P)	M	p.Leu44Val	-0.9	-1.4	-1.5	18.8
2.III.6	M	p.Leu44Val	-1.1	-1.3	-	-
2.II.2	M	p.Leu44Val	-	-0.8	-0.8	20.5
2.III.1	M	p.Leu44Val	0	-1.4	-	-
3.II.2	M	p.Leu44Val	-	-1.0	-0.9	18.6
3.III.2 (P)	M	p.Leu44Val	-0.9	-1.1	-1.3	18.7
3.II.3	M	p.Leu44Val		-0.4	-0.1	22.7
3.III.3	M	p.Leu44Val	-0.1	0.2	0.5	17.8

Table 1 – Clinical data of probands with FTO variants

Height is expressed in s.d. score (SDS) for national reference data for Finland at 4 years of age and at either 8 years for girls or 9 years for boys. Normal limits: delta HSDS <1.21, distance to target height at 4 yrs <1.76, distance to target height at 8/9 yrs <1.72(22). P – proband.



748

749

750 **Figure 1 – Flowchart of WES (whole exome sequencing) filtering**

751 **strategy to identify candidate genes.**

752

753

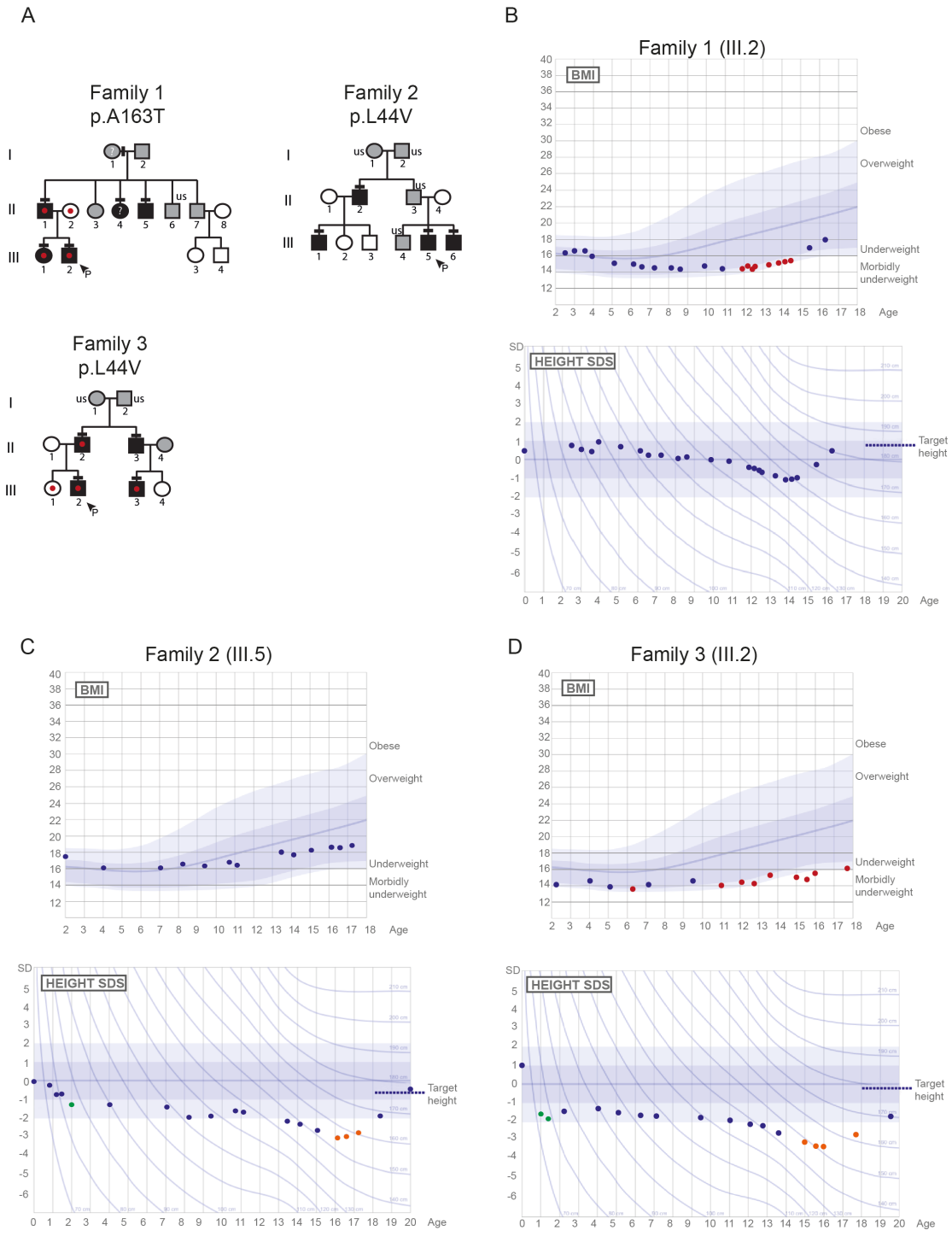
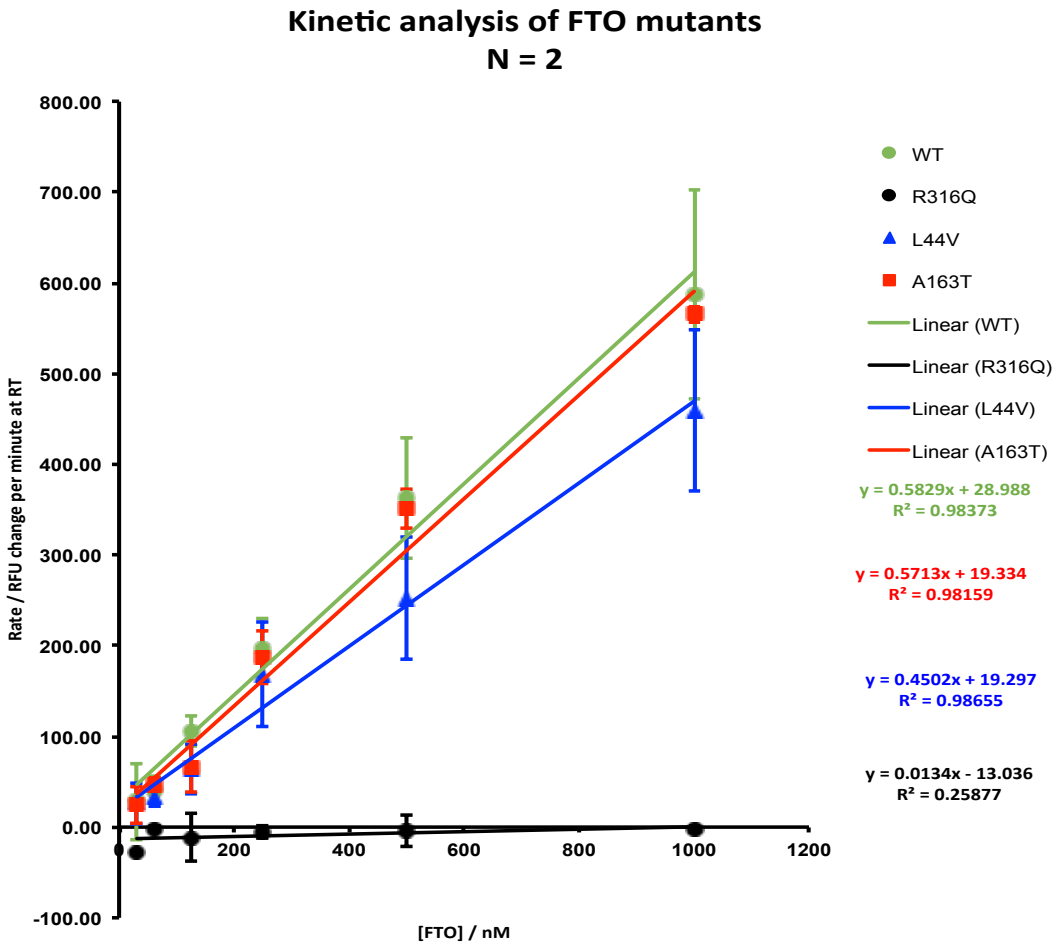


Figure 2 – Pedigrees and auxological data of the families with potentially pathogenic *FTO* variants



761

762

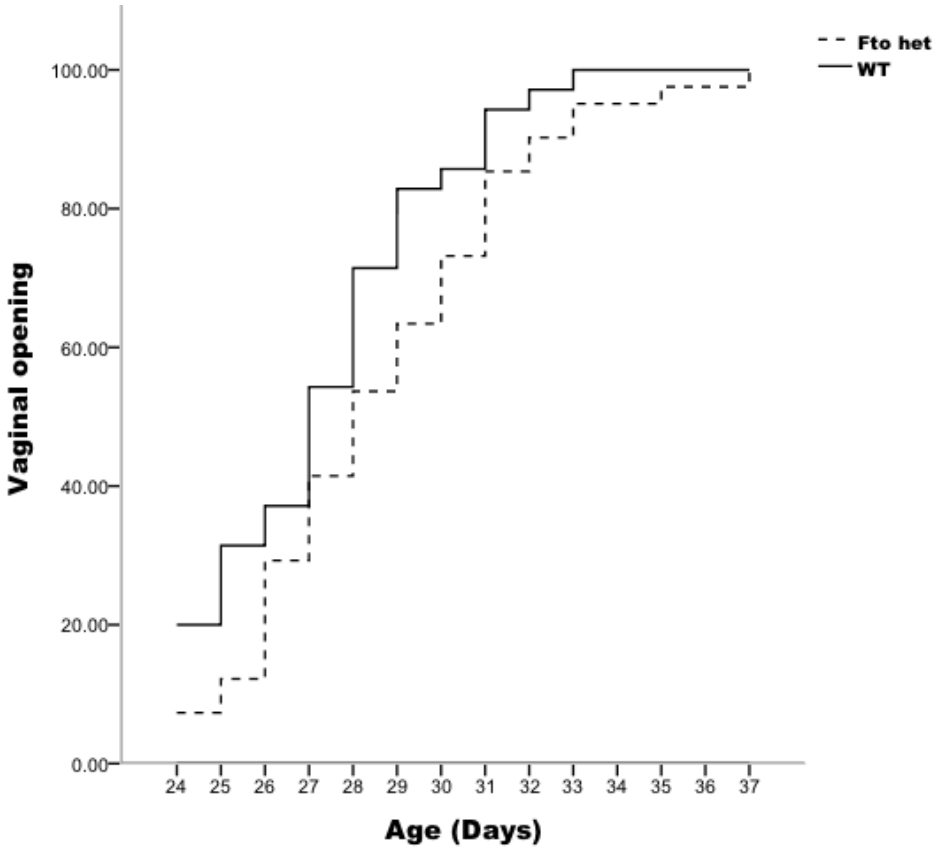
763 **Figure 3 – Demethylation assay assessing kinetic activity of mutant**

764 **versus wild type FTO proteins.**

765

766

767



768

769 **Figure 4 – Timing of vaginal opening in wild-type (WT) and *FTO*$^{+/-}$**

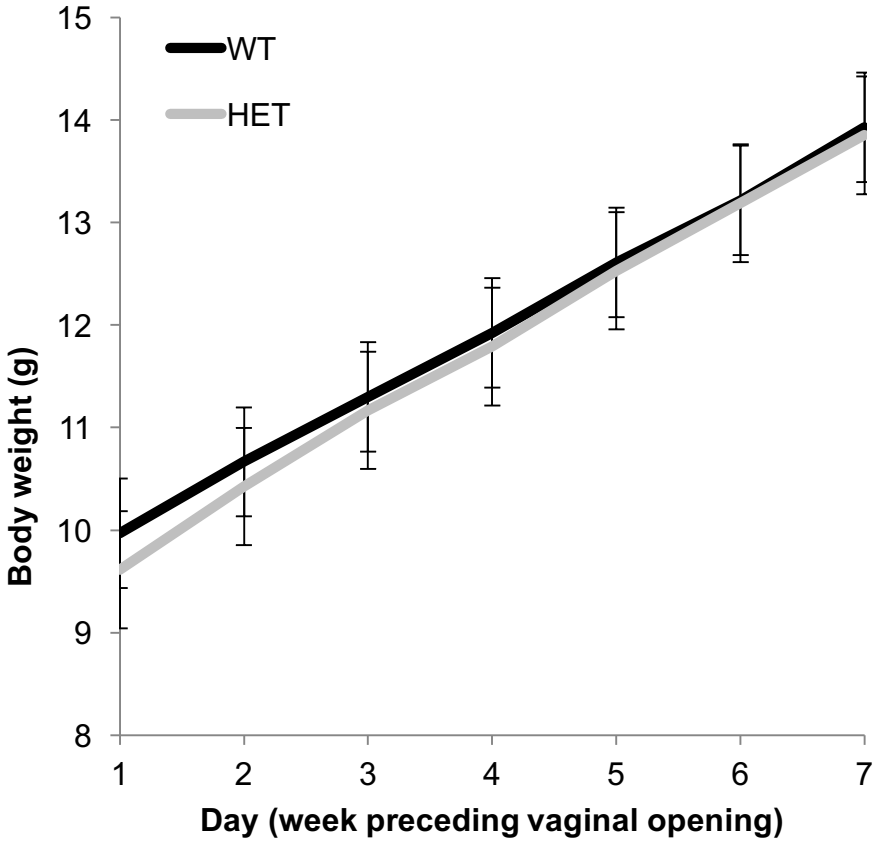
770 **heterozygous (Het) mice.**

771

772

773

774



775

776

777 **Figure 5 – Mean body weight (g) for wild type (WT) and *Fto*$^{+/-}$ (Het) mice in**

778 **7 days prior to vaginal opening**

779

Supplementary Data

Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Sasha R Howard, Leonardo Guasti, Ariel Poliandri, Alessia David, Claudia P Cabrera, Michael R Barnes, Karoliina Wehkalampi, Stephen O'Rahilly, Catherine E Aiken, Anthony P Coll, Marcella Ma, Debra Rimmington, Giles SH Yeo, Leo Dunkel

Table of Contents

Supplementary Methods:

Patient Details

For family members diagnosis of DP was based on PHV occurring 1.5 SD beyond the mean, i.e. age at takeoff exceeding 12.9 and 11.3 yr, or age at PHV exceeding 14.8 and 12.8 yr in males and females. This 1.5SD cutoff has sensitivity of 98% in identifying boys with Tanner stages G2 later than 14.0 years and sensitivity of 97% in identifying girls with Tanner stage B2 later than 13.0 years.

Genetic Analysis

Genetic analysis was performed in 160 individuals from the 67 most extensive families from our cohort with DP. These included 67 probands (male n=57, female n=10), 58 affected family members (male n=36, female n=22) and 35 unaffected family members (male, n=13, female n=22). Whole exome sequencing (WES) was performed on DNA extracted from peripheral blood leukocytes, using a Nimblegen V2 or Agilent V5 platform and Illumina HiSeq 2000 sequencing. The exome sequences were aligned to the UCSC hg19 reference genome. Picard tools and the genome analysis toolkit were used to mark PCR duplicates, realign around indels, recalibrate quality scores and call variants.

Variants were analyzed and filtered for potential causal variants in Ingenuity Variant Analysis (Qiagen) using filters for quality control, predicted functional annotation, minor allele frequency (MAF), and GWAS relevance (Figure 1). Quality control included thresholds for call quality, read depth and upstream pipeline filtering. Predicted functional annotation involved prioritizing nonsense, exonic missense, splice site variants, structural or promoter

changes, or variants deleterious to a microRNA. Filtering by MAF entailed including those variants with minor allele frequency (MAF) <1% in the 1000 Genomes database, the NHLBI exome variant server, and ExAC and gnoMAD databases. GWAS relevance filtering allowed identification of those remaining variants that lay within genes in linkage disequilibrium with 106 GWAS loci associated with AAM[1]. All genes in linkage disequilibrium with these GWAS AAM loci (using inclusive limits: D' > 0.8; r^2: no limit) were selected using the Broad institute SNAP tool (SNP annotation and proxy search). Linkage disequilibrium data was calculated using Haploview 4.0, based on phased genotype data from the International HapMap Project and the 1000 Genomes Project. A total of 760 genes were selected using this SNAP tool, and 'GWAS relevance filtering' allowed identification of those remaining variants that lay within these 760 genes[2]. Filters for genes implicated in body mass regulation were applied using a biological context filter with pathway analysis. Variants were then filtered for segregation with trait: variants present in ≥ n-1 affected individuals (where n = number of affected individuals in a given pedigree) and not present in more than one unaffected individual being retained. Family members were screened using conventional Sanger sequencing.

Targeted exome sequencing using a Fluidigm array of the remaining candidate gene identified post-filtering was then performed in a further 42 families from our cohort (288 individuals, 178 with DP; male=106, female=69 and 110 controls; male=55, female=58, Figure 1). Variants post targeted re-sequencing were filtered using the same criteria as the WES data: quality

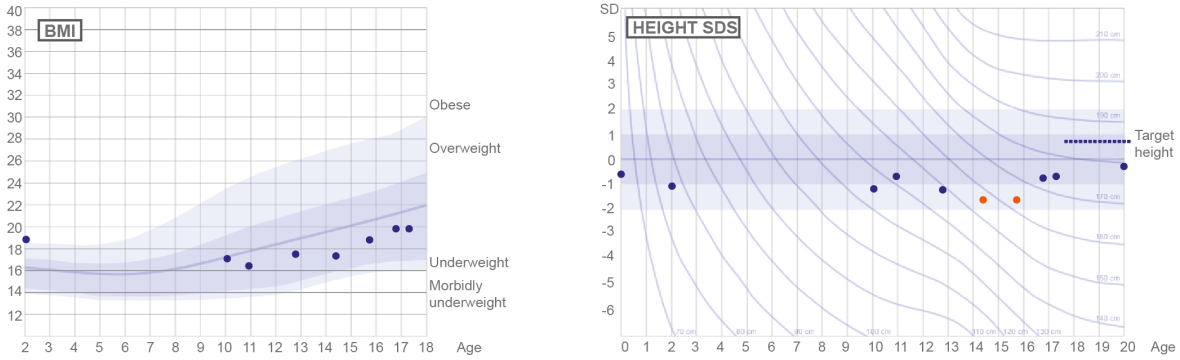
control, predicted functional annotation, minor allele frequency and segregation with trait.

Whole gene rare variant burden testing was performed post sequencing. Fisher's exact test was used to compare the prevalence of deleterious variants in our cohort with the Finnish population, using the ExAC Browser (Exome Aggregation Consortium (ExAC), Cambridge, MA: http://exac.broadinstitute.org, accessed September 2015). All variants from the ExAC database with minor allele frequency <1%, predicted to be deleterious by Polyphen-2 [3] or SIFT [4], were included in the analysis. A multiple comparison adjustment was applied post hoc using the Benjamini & Hochberg method [5], as detailed in [6]. Variants were confirmed via Sanger sequencing.

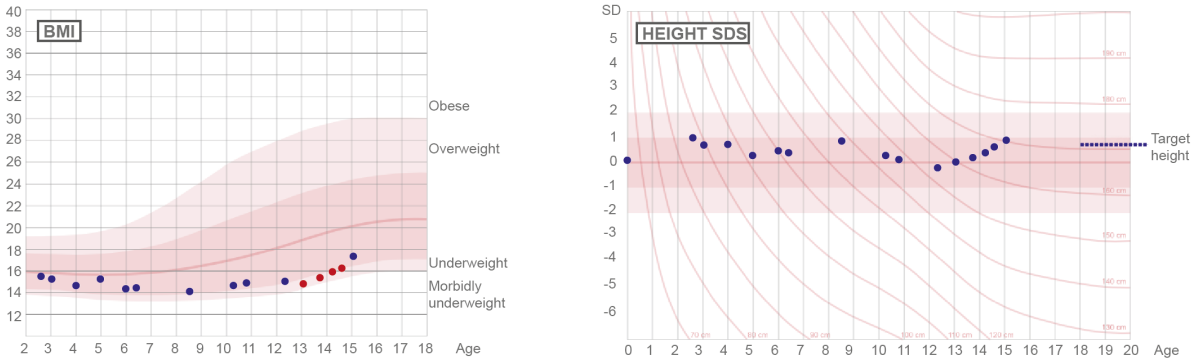
Supplementary Figures

Family 1 (p.A163T)

Family 1 (II.5)



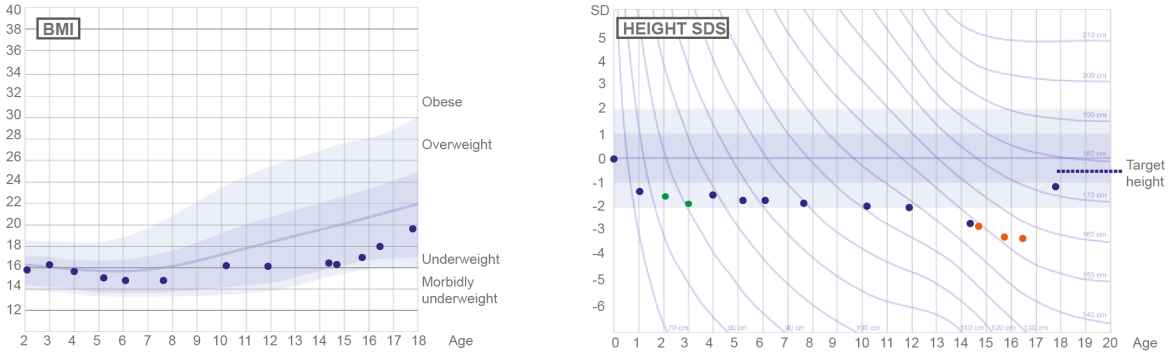
Family 1 (III.1)



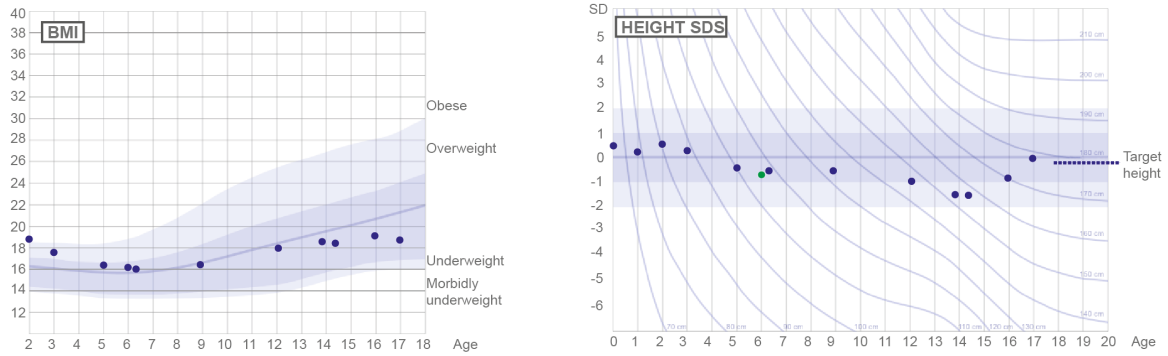
Supplementary Fig. 1 - Auxological data of the family members from Family 1 with potentially pathogenic *FTO* variant.
BMI and height standard deviation score (SDS) charts for the family members with *FTO* variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published [7].

Family 2 (p.L44V)

Family 2 (III.6)



Family 2 (II.2)



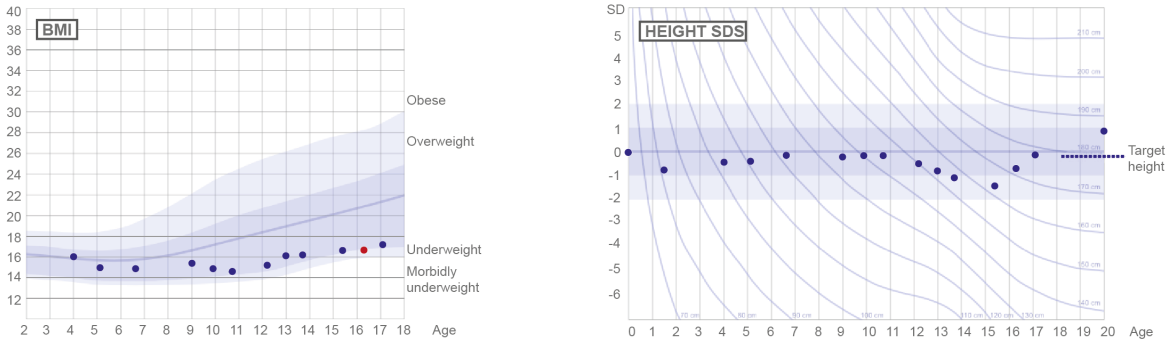
Supplementary Fig. 2 - Auxological data of the family members from Family 2 with potentially pathogenic *FTO* variant.
BMI and height standard deviation score (SDS) charts for the family members with *FTO* variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published [7].

Family 3 (p.L44V)

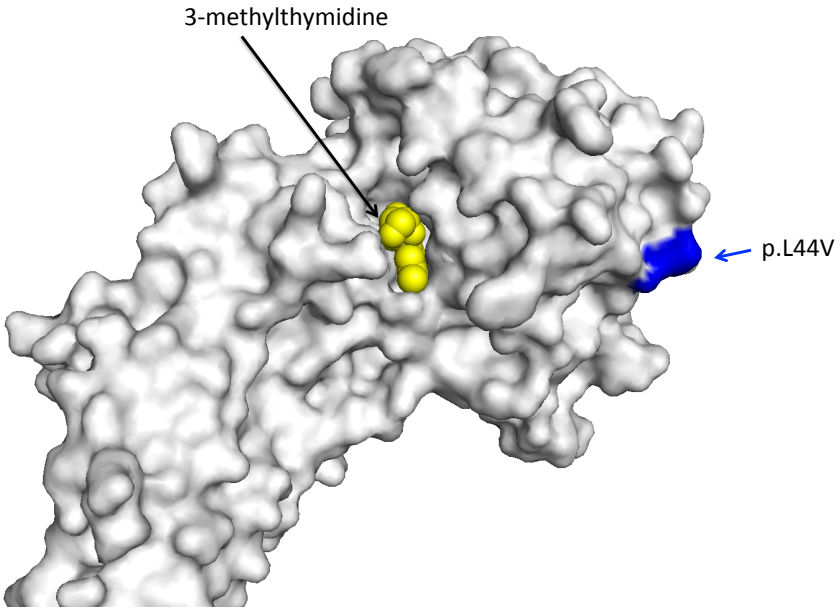
Family 3 (II.2)

Family 3 (II.3)

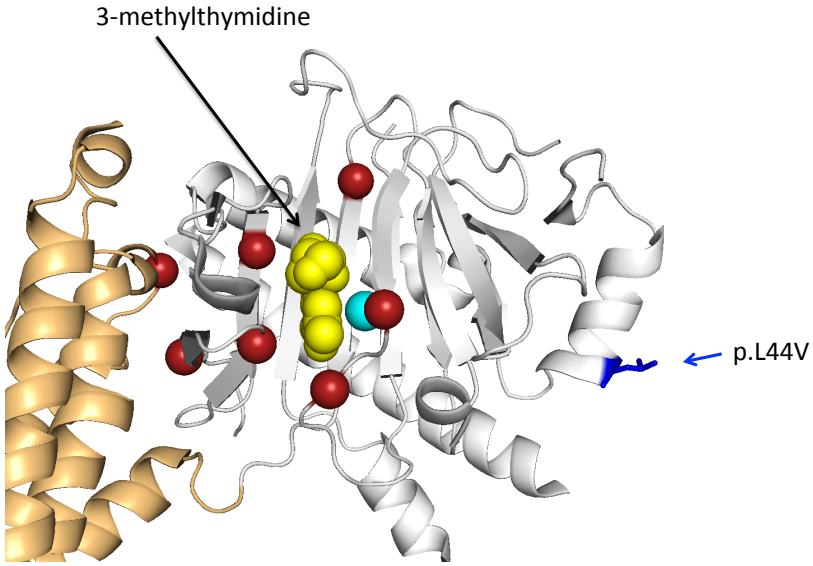
Family 3 (III.3)



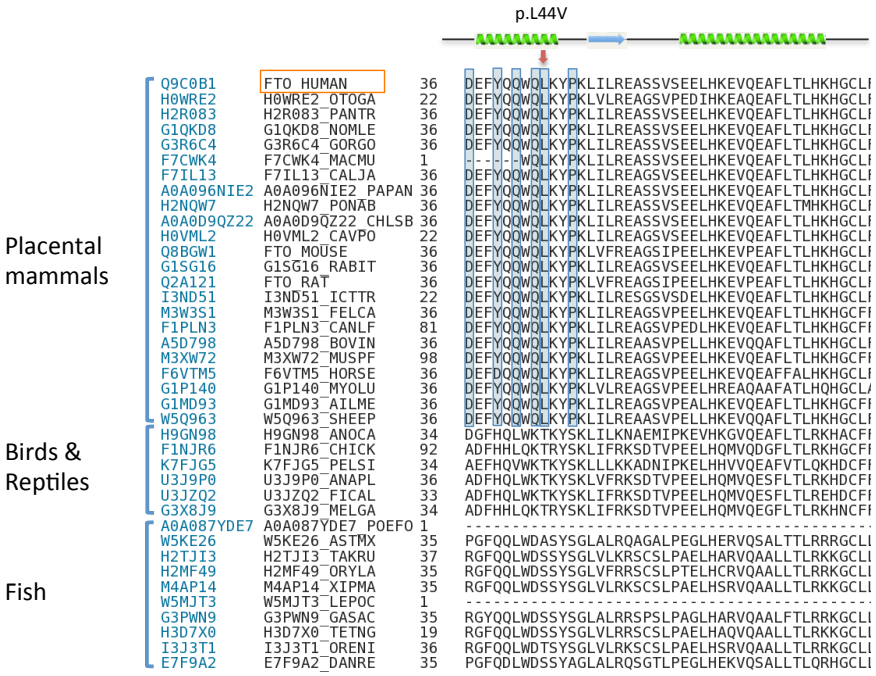
Supplementary Fig. 3 - Auxological data of the family members from Family 3 with potentially pathogenic *FTO* variant.
BMI and height standard deviation score (SDS) charts for the family members with *FTO* variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published [7].



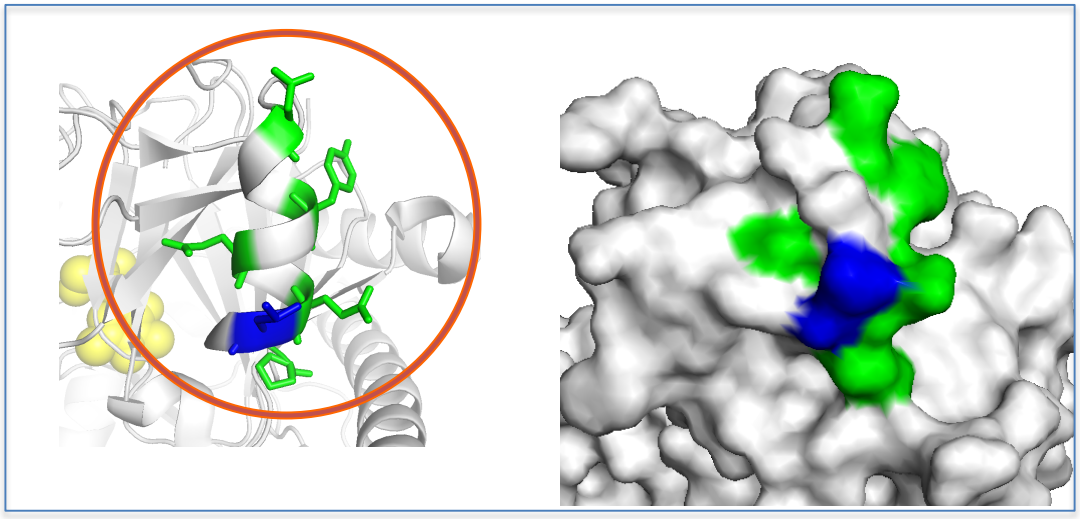
Supplementary Fig. 4 - Surface representation of FTO (here presented in grey) bound to 3---methylthymidine (in yellow). Leucine 44 is presented in blue.



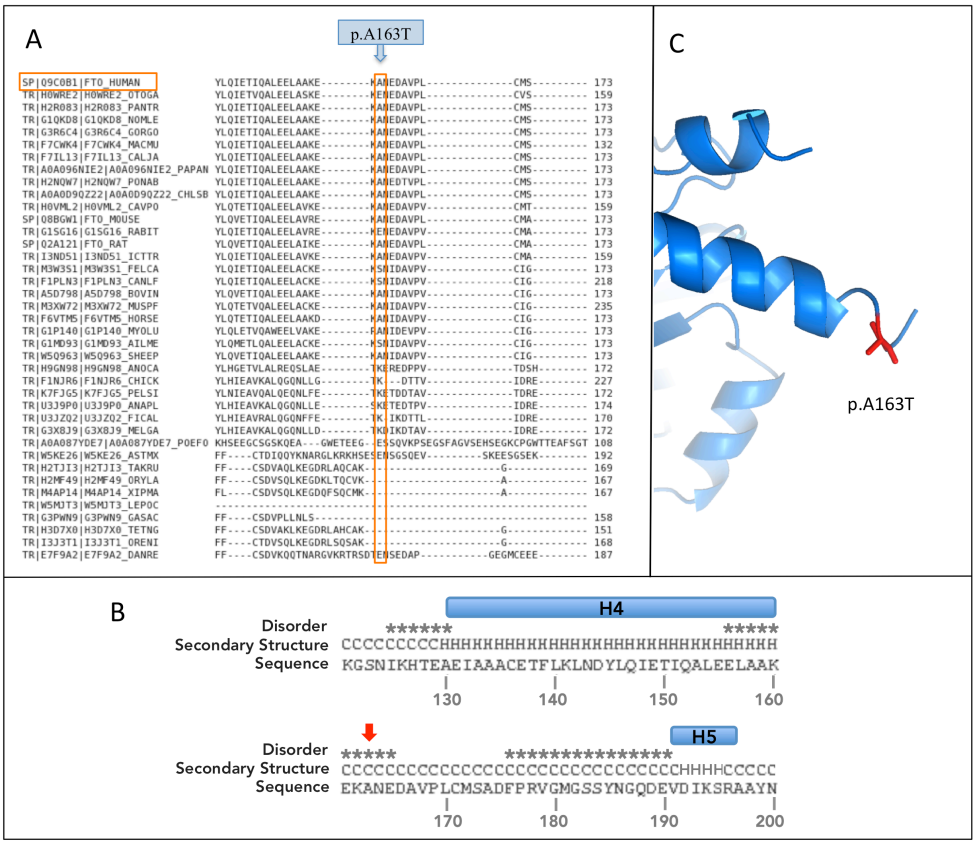
Supplementary Fig. 5 - 3D structure of FTO bound to 3---methylthymidine and iron (PDB 3flm)
FTO N---terminal domain is presented in grey and the C---terminal domain in orange. Dna is presented as yellow spheres, iron as a blue sphere. Deleterious mutations p.R96M or W, p. Y108A, p.F114D, p.E234P, p.C392D are presented as red spheres. p.L44V is presented in blue.



Supplementary Fig. 6 – Multiple sequence alignment between human FTO and its orthologues.
Orthologous sequences from other species were retrieved from Ensembl. Species are classified in “Placental mammals”, “Birds & Reptiles” and “Fish” according to Ensembl classification. For each species, FTO Uniprot accession number (in blue) and entry name (in black) are presented at the beginning of the row. The position of deleterious mutations on the human FTO sequence is indicated by a red arrow. L44 and other surrounding residues part of the same alpha helix form a motif, which is highly conserved across placental mammals but not in reptiles, birds and fish.



Supplementary Fig. 7 – Tertiary structure of FTO local to L44 residue. Residue L44 is presented in blue and other residues located on the same alpha helix and conserved across placental mammals are presented in green. FTO structure is shown as cartoon on the left and as surface on the right. 3---methylthymidine is presented as yellow spheres in the cartoon representation.



Supplementary Fig. 8 - p.A163T sequence and structural analysis.
Panel A: a multiple sequence alignment shows that alanine at position 163 is not conserved. Panel B: the position of Ala163 (indicated by a red arrow) is presented in relation to FTO secondary structure (H, alpha helix; C, residues that are not part of an alpha helix or a beta strand). Residues predicted to be disordered are indicated with an asterisk. Panel C: FTO tertiary structure. Ala163 (presented in red) is at the end of FTO alpha helix H4 and is not predicted to disrupt FTO function. FTO structure is visualized using visualisation program (http://www.pymol.org/).

Gene Symbol		
FTO	*GNPDA2*	*LEKR1*
SEC16B	*BDNF*	*DLK1*
TMEM18	*LIN28B*	*LEPR-LEPROT*
NEGR1	*SIX6*	
TNN13K	*CENPW-NCO7*	

Supplementary Table 1 – Genes involved in energy metabolism and growth pathways implicated in the timing of puberty in the general population from genome wide association studies (adapted from Perry et al [1]).

References

1. Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E, Ang WQ, Corre T, Cousminer DL, Feenstra B, Franceschini N, Ganna A, Johnson AD, Kjellqvist S, Lunetta KL, McMahon G, Nolte IM, Paternoster L, Porcu E, Smith AV, Stolk L, Teumer A, Tsernikova N, Tikkanen E, Ulivi S, Wagner EK, Amin N, Bierut LJ, Byrne EM, Hottenga JJ, Koller DL, Mangino M, Pers TH, Yerges-Armstrong LM, Hua Zhao J, Andrulis IL, Anton-Culver H, Atsma F, Bandinelli S, Beckmann MW, Benitez J, Blomqvist C, Bojesen SE, Bolla MK, Bonanni B, Brauch H, Brenner H, Buring JE, Chang-Claude J, Chanock S, Chen J, Chenevix-Trench G, Collee JM, Couch FJ, Couper D, Coviello AD, Cox A, Czene K, D'Adamo A P, Davey Smith G, De Vivo I, Demerath EW, Dennis J, Devilee P, Dieffenbach AK, Dunning AM, Eiriksdottir G, Eriksson JG, Fasching PA, Ferrucci L, Flesch-Janys D, Flyger H, Foroud T, Franke L, Garcia ME, Garcia-Closas M, Geller F, de Geus EE, Giles GG, Gudbjartsson DF, Gudnason V, Guenel P, Guo S, Hall P, Hamann U, Haring R, Hartman CA, Heath AC, Hofman A, Hooning MJ, Hopper JL, Hu FB, Hunter DJ, Karasik D, Kiel DP, Knight JA, Kosma VM, Kutalik Z, Lai S, Lambrechts D, Lindblom A, Magi R, Magnusson PK, Mannermaa A, Martin NG, Masson G, McArdle PF, McArdle WL, Melbye M, Michailidou K, Mihailov E, Milani L, Milne RL, Nevanlinna H, Neven P, Nohr EA, Oldehinkel AJ, Oostra BA, Palotie A, Peacock M, Pedersen NL, Peterlongo P, Peto J, Pharoah PD, Postma DS, Pouta A, Pylkas K, Radice P, Ring S, Rivadeneira F, Robino A, Rose LM, Rudolph A, Salomaa V, Sanna S, Schlessinger D, Schmidt MK, Southey MC, Sovio U, Stampfer MJ, Stockl D, Storniolo AM, Timpson NJ, Tyrer J, Visser JA, Vollenweider P, Volzke H, Waeber G, Waldenberger M, Wallaschofski H, Wang Q, Willemsen G, Winqvist R, Wolffenbuttel BH, Wright MJ, Australian Ovarian Cancer S, Network G, kConFab, LifeLines Cohort S, InterAct C, Early Growth Genetics C, Boomsma DI, Econs MJ, Khaw KT, Loos RJ, McCarthy MI, Montgomery GW, Rice JP, Streeten EA, Thorsteinsdottir U, van Duijn CM, Alizadeh BZ, Bergmann S, Boerwinkle E, Boyd HA, Crisponi L, Gasparini P, Gieger C, Harris TB, Ingelsson E, Jarvelin MR, Kraft P, Lawlor D, Metspalu A, Pennell CE, Ridker PM, Snieder H, Sorensen TI, Spector TD, Strachan DP, Uitterlinden AG, Wareham NJ, Widen E, Zygmunt M, Murray A, Easton DF, Stefansson K, Murabito JM & Ong KK. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* 2014 **514** 92-97.
2. Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, Lunetta KL, Visser JA, Byrne EM, Cousminer DL, Gudbjartsson DF, Esko T, Feenstra B, Hottenga JJ, Koller DL, Kutalik Z, Lin P, Mangino M, Marongiu M, McArdle PF, Smith AV, Stolk L, van Wingerden SH, Zhao JH, Albrecht E, Corre T, Ingelsson E, Hayward C, Magnusson PK, Smith EN, Ulivi S, Warrington NM, Zgaga L, Alavere H, Amin N, Aspelund T, Bandinelli S, Barroso I, Berenson GS, Bergmann S, Blackburn H, Boerwinkle E, Buring JE, Busonero F, Campbell H, Chanock SJ, Chen W, Cornelis MC, Couper D, Coviello AD, d'Adamo P, de Faire U, de Geus EJ, Deloukas P, Doring A, Smith GD, Easton DF, Eiriksdottir G, Emilsson V, Eriksson J, Ferrucci L, Folsom AR, Foroud T, Garcia M, Gasparini P, Geller F, Gieger C, Gudnason V, Hall P, Hankinson

SE, Ferreli L, Heath AC, Hernandez DG, Hofman A, Hu FB, Illig T, Jarvelin MR, Johnson AD, Karasik D, Khaw KT, Kiel DP, Kilpelainen TO, Kolcic I, Kraft P, Launer LJ, Laven JS, Li S, Liu J, Levy D, Martin NG, McArdle WL, Melbye M, Mooser V, Murray JC, Murray SS, Nalls MA, Navarro P, Nelis M, Ness AR, Northstone K, Oostra BA, Peacock M, Palmer LJ, Palotie A, Pare G, Parker AN, Pedersen NL, Peltonen L, Pennell CE, Pharoah P, Polasek O, Plump AS, Pouta A, Porcu E, Rafnar T, Rice JP, Ring SM, Rivadeneira F, Rudan I, Sala C, Salomaa V, Sanna S, Schlessinger D, Schork NJ, Scuteri A, Segre AV, Shuldiner AR, Soranzo N, Sovio U, Srinivasan SR, Strachan DP, Tammesoo ML, Tikkanen E, Toniolo D, Tsui K, Tryggvadottir L, Tyrer J, Uda M, van Dam RM, van Meurs JB, Vollenweider P, Waeber G, Wareham NJ, Waterworth DM, Weedon MN, Wichmann HE, Willemsen G, Wilson JF, Wright AF, Young L, Zhai G, Zhuang WV, Bierut LJ, Boomsma DI, Boyd HA, Crisponi L, Demerath EW, van Duijn CM, Econs MJ, Harris TB, Hunter DJ, Loos RJ, Metspalu A, Montgomery GW, Ridker PM, Spector TD, Streeten EA, Stefansson K, Thorsteinsdottir U, Uitterlinden AG, Widen E, Murabito JM, Ong KK & Murray A. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet* 2010 **42** 1077-1085.

3. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS & Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010 **7** 248-249.
4. Kumar P, Henikoff S & Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009 **4** 1073-1081.
5. Benjamini Y, Drai D, Elmer G, Kafkafi N & Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001 **125** 279-284.
6. Howard SR, Guasti L, Ruiz-Babot G, Mancini A, David A, Storr HL, Metherell LA, Sternberg MJ, Cabrera CP, Warren HR, Barnes MR, Quinton R, de Roux N, Young J, Guiochon-Mantel A, Wehkalampi K, Andre V, Gothilf Y, Cariboni A & Dunkel L. IGSF10 mutations dysregulate gonadotropin-releasing hormone neuronal migration resulting in delayed puberty. *EMBO Mol Med* 2016.
7. Saari A, Harju S, Makitie O, Saha MT, Dunkel L & Sankilampi U. Systematic growth monitoring for the early detection of celiac disease in children. *JAMA Pediatr* 2015 **169** e1525.